The Axiom of Choice, Zorn's Lemma and Well Ordering Principle

Abraham Si-Yiu Tung

10 Jun 2026 (Compiled July 30, 2026)

Contents

1 Introduction **2**

2 The proof to the equivalence **2**
2.1 AC $\implies$ ZL . . . 2
2.2 ZL $\implies$ WOP . . . 3
2.3 WOP $\implies$ AC . . . 4

3 Another proof on the equivalence: The other way around **4**
3.1 AC $\implies$ WOP . . . 4
3.2 WOP $\implies$ ZL . . . 4
3.3 ZL $\implies$ AC . . . 5

4 Some consequences of them **5**
4.1 (WOP) Sets can be associated with ordinals in a way that simplify proofs . . . 5
4.2 (AC) There will be subsets of $\mathbb{R}$ that have no measure; Banach-Tarski Theorem . . . 7
4.3 (AC) Every surjection has a right inverse . . . 8
4.4 (ZL) Each linear space have a basis
each ring has a maximal ideal
each field has an algebraic closure . . . 9

5 What if they don't hold? **10**
5.1 ($\neg$WOP) There will be sets that are *not* equinumerous with any ordinal . . . 10
5.2 ($\neg$AC) There will be two sets whose cardinalities cannot be compared . . . 11
5.3 ($\neg$ZL) There will be a linear space that has no basis . . . 11

1 Introduction

This article proves the equivalence of the Axiom of Choice, Zorn's Lemma, and the Well-Ordering Principle under ZF, and investigate some of the typical results when the axioms hold or fail.

This is the *Axiom of Choice* (AC).

Axiom 1.1. *If F is a set of nonempty sets, then there is a mapping $h : F \to \bigcup F$ such that for each $f \in F$, $h(f) \in f$.*

This is the *Zorn's Lemma* (ZL).

Axiom 1.2. *If P is a* nonempty *set such that*

- *P is partially ordered by $\leq$.*
- *For each subset T of P, if T is linearly ordered by $\leq$, then T has an upper bound in P.*

Then P has at least one maximal element, i.e. an element $m \in P$ such that

$$\forall p \in P(p \geq m \implies p = m).$$

And this is the *Well Order Principle/Theorem* (WOP).

Axiom 1.3. *If P is a set, then we can define a well order on it.*

2 The proof to the equivalence

This section will proof the equivalence of AC, ZL and WOP by proving AC $\implies$ ZL $\implies$ WOP $\implies$ AC.

2.1 AC $\implies$ ZL

Theorem 2.1. AC $\implies$ ZL

Proof. Assume for the sake of contradiction that P is a set such that

- P is partially ordered by $\leq$;
- every chain (subset linearly ordered by $\leq$) of P has an upper bound in P;
- P does *not* have a maximal element.

Intuition might suggest that every maximal element is also the largest element of some chains. Hence we can look for a chain that goes on indefinitely. We can perform the following operations:

1. Take $a_0 \in P$

2. Let α be an ordinal. If a_β has been defined for all ordinals $\beta < \alpha$, and $(a_\beta)_{\beta<\alpha}$ is a chain, then let a_α be a *strict* upper bound of $(a_\beta)_{\beta<\alpha}$.

Then a_α can be defined for all ordinals α, which is a contradiction, because the chain (which is a subset of P) will be too large, and contradicts the Hartog's Number Theorem. Hence P must have a maximal element. □

One might argue that AC is not really involved in this proof, but this is not the case: We can formalize the choice of a_α. Let K be the set of all chains, and for $C \in K$ let

$$U(C) := \{p \in P : p \text{ is a } \textit{strict} \text{ upper bound of } C\},$$

then $U(C) \neq \emptyset$ for each $C \in K$.

Then we define

$$h : \{U(C) : C \in K\} \to P$$

such that for each $C \in K$, $h(U(C)) \in U(C)$. This step is used in defining a_α when a_β is defined for all ordinals $\beta < \alpha$, and this step relies on AC.

Besides the subtle use of AC in the proof, another thing worth mentioning is the methodology: We have used *transfinite induction* to define a_α.

2.2 ZL $\implies$ WOP

Theorem 2.2. ZL $\implies$ WOP

Proof. Assume ZL holds. We let

$$\Omega := \{R \subseteq P \times P : R \text{ is a well-ordering on a subset of } P\}$$

Since we can define well order on all finite subsets of P, Ω is nonempty as long as P is nonempty. For each $R \in \Omega$, let

$$\mathrm{Dom}(R) := \{w \in P : \exists z \in P((w, z) \in R \vee (z, w) \in R)\}$$

Then we define a order on Ω: If $R, S \in \Omega$, then $R \preceq S$ if and only if $\mathrm{Dom}(R) \subseteq \mathrm{Dom}(S)$, $R = S \cap (\mathrm{Dom}(R) \times \mathrm{Dom}(R))$, and $\mathrm{Dom}(R)$ is an initial segment of $\mathrm{Dom}(S)$. Then $\preceq$ is a partial ordering on Ω. Also, if T is a chain in Ω, then $\bigcup T$ is an upper bound of T in Ω. According to ZL, Ω has a maximal element S.

If $\mathrm{Dom}(S) \neq P$, then let $p \in P \setminus \mathrm{Dom}(S)$ and

$$S' := S \cup ((\mathrm{Dom}(S) \cup \{p\}) \times \{p\})$$

then $S \subsetneq S'$ and $S' \in \Omega$, contradicting with the maximality of S. Hence S is a well ordering on P. □

2.3 WOP $\implies$ AC

Theorem 2.3. WOP $\implies$ AC

Proof. Assume that WOP holds. We can well order $\bigcup F$, then let the choice function h map each $f \in F$ into its minimal element. □

3 Another proof on the equivalence: The other way around

This section proves AC $\implies$ WOP $\implies$ ZL $\implies$ AC, which is another valid proof for the equivalence of AC, ZL and WOP.

3.1 AC $\implies$ WOP

Theorem 3.1. AC $\implies$ WOP

Proof. Let P be the set on which one wants to find a well-order. Then there is a choice function $h : 2^P \setminus \{\emptyset\} \to P$.

We may take elements a_α from P via the transfinite inductive formula below:

$$a_\alpha := h(P \setminus \{a_\beta : \beta \in \alpha\})$$

where $\alpha \in \mathbf{On}$ such that $\{a_\beta : \beta \in \alpha\} \neq P$. Then there is an ordinal α_0 such that

- a_α is defined for all $\alpha \in \alpha_0$, and
- a_{α_0} is *not* defined.

Hence $(a_\alpha)_{\alpha \in \alpha_0}$ is a well-ordering of P. □

3.2 WOP $\implies$ ZL

Theorem 3.2. WOP $\implies$ ZL

Proof. Let P be a partially ordered set on which every chain has an upper bound. Assume for the sake of contradiction that P has no maximal element. Then for each $p \in P$, there exists $p' \in P$ such that $p < p'$.

From WOP, we can find a well-ordering R of P. For each $S \in P$, let $\min S$ denote the R-minimal element of S. Then use transfinite induction to define

$$a_\alpha = \min \{s \in P : s \text{ is a } \textit{strict} \leq\text{-upper bound of } \{a_\beta : \beta \in \alpha\}\}$$

for $\alpha \in \mathbf{On}$. Since no element of P is $\leq$-maximal, a_α can be defined for each $\alpha \in \mathbf{On}$. Then every ordinal injects into P, which contradicts with the Hartog's Number Theorem. □

3.3 ZL $\Longrightarrow$ AC

Theorem 3.3. ZL $\Longrightarrow$ AC

Proof. Let C be the set of all choice functions on the subsets of F. C is nonempty, since a choice function can always be defined on a finite set. Then C is partially ordered by $\subseteq$. Also, if H is a chain in C, then $\bigcup H$ is its upper bound in C. Hence from ZL, C has a maximal element h. If h is not defined on some nonempty subset of F, we can always extend h, which contradicts the maximality of h. Therefore h is defined on all nonempty subsets of F. □

4 Some consequences of them

If we assume AC, ZL and WOP hold, then these are some typical consequences.

4.1 (WOP) Sets can be associated with ordinals in a way that simplify proofs

First we introduce some notations:

Definition 4.1. *If X and Y are sets, then*

- *$X \preceq Y$ means that there exists a one-one mapping from X to Y;*
- *$X \approx Y$ means that there exists a bijection from X to Y;*
- *$X \prec Y$ means $X \preceq Y \wedge \neg(X \approx Y)$.*

According to [1], the definition of *ordinals* and *cardinals* are as following:

Definition 4.2. *A set α is an* ordinal *if and only if*

- *α is well ordered by $\in$;*
- *All elements of α are subsets of α.*

All ordinals form a class, usually denoted with **On**. Note that **On** is a proper class (it is too large to be a set).

Theorem 4.3. **On** *is well ordered by $\subseteq$.*

Proof. $\subseteq$ is a partial order on any set. Hence it suffices to show

- (Each pair of elements is comparable) If α and β are ordinals, then either $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.
- (Every subset has a minimal element) If S is a set of ordinals, then S contains a minimal element.

See lemma 1.2.9 and theorem 1.2.10 of [1] for a proof. □

Definition 4.4. *An ordinal α is a* cardinal *if and only if*

$$\forall \beta < \alpha (\beta \not\approx \alpha).$$

Theorem 4.5. *If A is a well ordered set, then there is exactly one ordinal α such that*

$$A \cong \alpha.$$

Proof.

- We show the existence of such α first.

  We can define $\phi : A \to \mathbf{On}$ such that

  $$\phi(a) = \{\phi(b) : b \in A \wedge b < a\}$$

  for all $a \in A$. It suffices to show $\phi(A)$ is an ordinal.

  - $\phi(A)$ is well ordered by $\in$: This can be deduced directly from

    $$\phi(a) = \{\phi(b) : b \in A \wedge b < a\}.$$

  - Each element of $\phi(A)$ has format

    $$\{\phi(b) : b \in A \wedge b < a\}$$

    and is thus a subset of $\phi(A)$.

  Therefore $\phi(A)$ is an ordinal. Now, ϕ is monotonically increasing, and for each $a \in A$, $\phi(a)$ is the *smallest* ordinal greater than each of $\phi(b)$ ($b \in A$), hence ϕ is an order-isomorphism from A to $\phi(A)$. Hence the desired conclusion is true.

- The uniqueness of α can be deduced from the fact that two different ordinals are not order-isomorphic.

□

Note. To be clear, this theorem does not depend on WOP. It merely depends on the ZF Axiom Schema of Replacement. But we list it here because WOP asserted that every set can be given a well-order. □

Theorem 4.6. *If X is a set, then there is a unique* cardinal *x such that a bijection can be built from X to x.*

There are other nice consequences.

Theorem 4.7. *If X and Y are sets, then*

1. *Exactly one of the three statements hold:*

   - $X \prec Y$;
   - $X \approx Y$;

- $X \succ Y$.

2. *(Cantor-Schröeder-Berinstein Theorem) If X and Y are sets and $X \preceq Y$, $Y \preceq X$, then $X \approx Y$.*

Proof. The reason for both of them is that the cardinals are well ordered by $\preceq$. □

Note. The trichotomy property is unprovable in ZF and requires WOP.

However, the Cantor-Schröeder-Bernstein Theorem merely requires ZF. We list the theorem here because the use of WOP can simplify its proof. □

4.2 (AC) There will be subsets of $\mathbb{R}$ that have no measure; Banach-Tarski Theorem

We all know that all of the intervals $[0,1]$, $[0,1)$, $(0,1]$ and $(0,1)$ have length 1. However, we may want to define a 'length' rigorously. Use $\mu : 2^{\mathbb{R}} \to \mathbb{R}_{\geq 0}$ be a function that evaluates length. Then we may want the following properties:

(*Empty Set*) $\mu(\emptyset) = 0$.

(*Normalization*) $\mu([0,1]) = 1$.

(*Positivity*) $0 \leq \mu(S) < +\infty$ for each $S \subseteq \mathbb{R}$.

(*Monotonicity*) If $A \subseteq B \subseteq \mathbb{R}$, then $\mu(A) \leq \mu(B)$.

(*Countable sub-additivity*) If $\{A_i\}_{i \in I}$ is an *at most countable* family of sets, then

$$\mu\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} \mu(A_i).$$

(*Countable additivity*) If $\{A_i\}_{i \in I}$ is an *at most countable* family of *disjoint* sets, then

$$\mu\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mu(A_i).$$

(*Translational Invariance*) If $A \subseteq \mathbb{R}$ and $x + A := \{x + a : a \in A\}$, then $\mu(A) = \mu(x + A)$.

The thing is: Using AC, we can find a subset of $\mathbb{R}$ that cannot have any length.

Theorem 4.8. *There exists $E \subseteq \mathbb{R}$ such that $\mu(E)$ cannot be defined.*

Proof. We use the axiom of choice. For each $q \in \mathbb{R}$, define

$$M(q) := \{p \in [0,1] : p - q \in \mathbb{Q}\}.$$

Then let F be the collection of these $M(q)$s. From AC, we can take an element from each $M(q)$, forming a new set E.

Assume for the sake of contradiction that $\mu(E)$ exists. We have

$$[0,1] \subseteq \bigcup_{p \in \mathbb{Q} \cap [-1,1]} E + p \subseteq [-1,2].$$

Since $\mathbb{Q}$ is countable, we obtain from countable additivity of measures that

$$1 \leq \sum_{p \in \mathbb{Q} \cap [-1,1]} \mu(E+p) \leq 3.$$

Due to the translational invariance, there is

$$1 \leq \sum_{p \in \mathbb{Q} \cap [-1,1]} \mu(E) \leq 3.$$

Then $\mu(E)$ cannot be a real number, contradiction.

Hence E has no measure. □

It can be shown that if AC does not hold and some other conditions hold, then we can build a $\mathbb{R}$ where each subset is measurable.

Another bizarre consequence of AC is the Banach-Tarski Theorem: We can disassemble a sphere of radius 1 into finitely many pieces, then reassemble the pieces into two spheres with radius 1.

Theorem 4.9. *(Banach-Tarski) If $\mathbb{S}$ is the unit sphere surface in $\mathbb{R}^3$, then we can find $n \in \mathbb{N}$ and subsets $\Omega_1, \Omega_2, \ldots, \Omega_n$ of $\mathbb{S}$ such that*

- $\mathbb{S} = \Omega_1 \sqcup \Omega_2 \sqcup \cdots \sqcup \Omega_n$;
- *There exist rigid motions $P_1, P_2, \ldots, P_n$ and $1 \leq p < n$ such that*

$$\begin{aligned} \mathbb{S} &= P_1(\Omega_1) \sqcup \cdots \sqcup P_p(\Omega_p) \\ &= P_{p+1}(\Omega_{p+1}) \sqcup \cdots \sqcup P_n(\Omega_n) \end{aligned}$$

It is worthwhile noting that the pieces we have divided $\mathbb{S}$ into are not measurable, so it is another way of proving the existence of non-measurable sets when AC holds.

4.3 (AC) Every surjection has a right inverse

Theorem 4.10. *If $f : X \to Y$ and $f(X) = Y$, then there is $g : Y \to X$ such that $f \circ g = \mathrm{Id}_Y$.*

Proof. For each $y \in Y$, we let

$$P(y) := \{x \in X : f(x) = y\}.$$

Then $\bigcup_{y \in Y} P(y) = X$ and $P(y) \neq \emptyset$ for each $y \in Y$. Hence we can find a choice function $h : \{P(y) : y \in Y\} \to X$ such that for each $y \in Y$, $h(P(y)) \in P(y)$. If we regard P as a mapping from Y to 2^X, then $h \circ P$ is the right inverse we are looking for. □

4.4 (ZL) Each linear space have a basis each ring has a maximal ideal each field has an algebraic closure

Theorem 4.11. *Every linear space V has a basis.*

This proof comes from the Proposition A.3.3, Appendix A.3 of [2].

Proof. We use ZL. Let

$$P := \{T \subseteq V : T \text{ is linear independent}\}$$

be a set partially ordered by $\subseteq$. Then for each linearly ordered subset T of P, $\bigcup T \in P$ and $\bigcup T$ is an upper bound of T. From ZL, we can find a maximal element M of P. Now, if $v \notin \text{Span}M$, then $M \sqcup \{v\}$ is a larger element of P, which contradicts with the maximality of M. Hence M is a basis of V. □

Theorem 4.12. *Every ring R has a maximal proper ideal.*

Proof. Let Z be the set of all *proper* ideals of R, partially ordered by $\subseteq$. Then for each linearly ordered subset T of Z, $\bigcup T$ is an upper bound of T in Z. Hence, according to ZL, Z has a maximal element, which is automatically a maximal ideal. □

Theorem 4.13. *Every field $\mathbb{F}$ has an algebraic closure.*

This proof comes from [5].

Proof. Let S be the set of all monic irreducible polynomials in $\mathbb{F}[x]$. For each $f_\lambda \in S$, let $u_{1,\lambda}, u_{2,\lambda}, \ldots, u_{d,\lambda}$ be roots of f_λ (d is the degree of f_λ). Let R be the polynomial ring over $\mathbb{F}$ generated by $u_{i,\lambda}$ for all $\lambda \in \Lambda$ and all $i \leq \deg(f_\lambda)$. Write

$$f_\lambda - \prod_{i=1}^{d} (x - u_{i,\lambda}) = \sum_{j=0}^{d-1} r_{j,\lambda} x^j \in R[x]$$

with $r_{i,\lambda} \in R$. Let I be the ideal in R generated by the $r_{i,\lambda}$. Since I is strictly smaller than R, theorem 4.12 implies that there exists a maximal ideal M in R that contains I. The field $K_1 = R/M$ has the property that every polynomial f_λ with coefficients in K splits as the product of $x - (u_{i,\lambda} + M)$, and hence has all roots in K_1. In the same way, an extension K_2 of K_1 can be constructed, etc. The union of the extensions K_m ($m \in \mathbb{N}$) is the algebraic closure of $\mathbb{F}$, because any polynomial with coefficients in this new field has its coefficients in some K_n with sufficiently large n, and then its roots are in K_{n+1}, and hence in the union of all K_ms.

It can be shown along the same lines that for any subset S of $\mathbb{F}[x]$, there exists a splitting field of S over K. □

5 What if they don't hold?

We explore the world without AC, ZL and WOP.

5.1 (¬WOP) There will be sets that are *not* equinumerous with any ordinal

Theorem 5.1. *If* ¬WOP *holds, then there is a set A that is not equinumerous with any ordinal.*

Proof. Let A be a set that cannot be well-ordered. Then for each ordinal α, there is no injection from A to α, because otherwise we can well-order A using this injection. Hence A is not equinumerous with any ordinal. □

However, this does not mean we can find a set that is larger than any ordinal.

1. If 'X is larger than Y' means there is an injection from Y to X, then we have the following theorem.

   Theorem 5.2. *(The Hartog's Number Theorem) For each set X, there is an ordinal $\aleph(X)$ (called the* Hartog's Number*) that does* not *inject into X.*

   This theorem is the theorem 8.18 of [3] and the proof is also available in the 'Proof' section of [4].

   Proof. Let X be any set. We may construct a class

   $$\alpha := \{\beta \in \mathbf{On} : \beta \preceq X\}.$$

   Intuitively, we might think that α is an ordinal, and we can prove it here.

   (a) α is a set. We can let

   $$Y := \{(A, R) : A \subseteq X, (A, R) \text{ is a well-ordered set}\}.$$

   Then we define a formula

   $$\phi(y, \gamma) = (y \cong \gamma).$$

   Then ϕ determines a function from Y to **On** due to theorem 4.5. Also, $\alpha = \mathrm{Cod}(\phi)$. This is because

   - For each $(A, R) \in Y$, $A \subseteq X$, hence $\phi(A) \preceq X$, and $\phi(A) \in \alpha$.
   - For each $\beta \preceq X$, there is a mapping $\iota : \beta \hookrightarrow X$, and also, β is well-ordered by $\in$, hence $\mathrm{Cod}(\iota) \subseteq X$ is well-ordered by $\in$. Hence $(\mathrm{Cod}(\iota), \in) \in Y$.

   From the ZF Axiom Schema of Replacement, α is a set.

(b) α is an ordinal. Since α is a set of ordinals, it is well-ordered by $\in$. Also, for each $\gamma \in \mathbf{On}$,

$$\gamma = \{\beta \in \mathbf{On} : \beta \prec \gamma\}.$$

Hence for each $\beta \in \alpha$, if $\gamma \in \beta$, then $\gamma \prec \beta \preceq X$, which implies $\gamma \in \alpha$. Hence each element of α is a subset of α, and α is an ordinal.

(c) If $\alpha \preceq X$, then $\alpha \in \alpha$, which contradicts with the ZF Axiom of Regularity. Hence $\alpha \not\preceq X$.

Hence α is an ordinal that is 'larger' than X. □

2. If 'X is larger than Y' means there is a surjection from X to Y, then we have the following theorem.

Theorem 5.3. *For each set X, there is an ordinal α that X does* not *surject into.*

Before the proof of theorem 5.3, we need a lemma.

Lemma 5.4. *(The Bruali-Forti Paradox)* **On** *is* not *a set.*

Proof of Lemma 5.4. Assume for the sake of contradiction that **On** is a set. Then since **On** is well-ordered by $\in$, and each element of **On** (each ordinal) is a subset of **On**, we have $\mathbf{On} \in \mathbf{On}$, which contradicts with the ZF Axiom of Regularity. □

Proof of Theorem 5.3. Assume for sake of contradiction that such X exists. Then there is a mapping $g : X \twoheadrightarrow \mathbf{On}$. This is a contradiction, because $\mathbf{On} = \mathrm{Cod}(g)$ is *not* a set according to lemma 5.4. □

This is a common pitfall: If WOP does not hold, then intuition may tell one that there exists a very large set. However it is not the case.

5.2 (¬AC) There will be two sets whose cardinalities cannot be compared

Theorem 5.5. *If ¬AC holds, then there are sets X and Y such that none of the three statements hold.*

- $X \prec Y$
- $X \approx Y$
- $X \succ Y$

5.3 (¬ZL) There will be a linear space that has no basis

Theorem 5.6. *If ¬ZL holds, then there is a linear space that has no basis.*

References

[1] 李文威《代数学方法-卷 I：基础架构》(available on https://github.com/wenweili/AlJabr-1)

[2] Michael Artin, *Algebra* (Second Edition).

[3] Derek Goldrei, *Classic Set Theory for Guided Independent Study*

[4] Wikipedia Page on Hartog's Number, https://en.wikipedia.org/wiki/Hartogs_number

[5] Wikipedia Page on Algebraic Closures, https://en.wikipedia.org/wiki/Algebraic_closure#Existence_of_an_algebraic_closure_and_splitting_fields